Proc. Nat. Acad. Sci. USA
Vol. 72, No. 5, pp. 1654–1657, May 1975

Some Distributions Associated with Bose–Einstein Statistics

(skew distributions/Pareto Law/city sizes)

YUJI IJIRI AND HERBERT A. SIMON

Carnegie-Mellon University, Pittsburgh, Pennsylvania 15213

Contributed by Herbert A. Simon, February 18, 1975

ABSTRACT **This paper examines a stochastic process for Bose–Einstein statistics that is based on Gibrat's Law (roughly: the probability of a new occurrence of an event is proportional to the number of times it has occurred previously). From the necessary conditions for the steady state of the process are derived, under two slightly different sets of boundary conditions, the geometric distribution and the Yule distribution, respectively. The latter derivation provides a simpler method than the one earlier proposed by Hill [*J. Amer. Statist. Ass.* (1974) 69, 1017–1026] for obtaining the Pareto Law (a limiting case of the Yule distribution) from Bose–Einstein statistics. The stochastic process is applied to the phenomena of city sizes and growth.**

It is well known (p. 61 of ref. 1) that under appropriate conditions the geometric distribution is a limiting distribution for Bose–Einstein statistics. Indeed, in applications to physics, it is usually treated as though it were the only limiting distribution. However, in a recent paper Hill (2) has shown that under different (and much more complicated) conditions, Bose–Einstein statistics can be made to yield the Pareto distribution as a limiting distribution. Since the Pareto distribution often gives an excellent fit to data on the relative frequencies of cities of different sizes, Hill has offered his derivation of this distribution from Bose–Einstein statistics as an "explanation" of the observed city size distributions.

We do not intend to examine formally here the meaning of "explanation." Nevertheless, to say that cities are distributed by size according to the Pareto Law *because* they obey Bose–Einstein statistics would appear, in common-sense terms, to be less an explanation than a relocation of the mystery. Cities change size as a result of the births and deaths that take place within them, of the migrations into them from non-urban or foreign places, and the migrations between pairs of them. A satisfying explanation of the observed size distributions in terms of Bose–Einstein statistics would have to show a relation between these statistics and the birth, death, and migration processes listed above.

There already exist in the literature several derivations of the Pareto Law from plausible stochastic assumptions about birth, death, and migration processes (3, 4). Although they differ in their details, all of these derivations have in common some variant of what is often called Gibrat's Law. This law, applied to cities, states that the accretions to the population of a city (or a group of cities of nearly equal population) by births and migration will occur at a rate (per capita) nearly independent of the present city size; and that the loss of inhabitants by deaths and migration will also occur at a rate nearly independent of city size. More generally, Gibrat's Law postulates expected growth proportional to size. Under this assumption, we have something like a random walk on a logarithmic scale, and would expect to derive from it highly skewed limiting distributions related to the log normal. Indeed, this is what we find, for by small changes in the boundary conditions and other parameters of our stochastic process we can obtain the log normal distribution, the Yule distribution and its limiting form, the Pareto distribution, Fisher's logarithmic series, and the negative binomial.

In this paper, we will show first that Bose–Einstein statistics satisfy Gibrat's Law. By this means, we will explain why the statistics of city size can be derived from Bose–Einstein statistics. Second, we will interpret the Bose–Einstein scheme in terms of stochastic processes to show what boundary conditions would lead to the geometric distribution, and what boundary conditions would lead to the Pareto distribution. Third, we will discuss the interpretations of the alternative assumptions in terms of city growth processes. Derivations of skewed distributions from Bose–Einstein statistics had apparently not been known until Hill's article appeared. Our derivation of the Pareto distribution from Bose–Einstein statistics will use much simpler methods, and much weaker assumptions, than those employed by Hill.

Bose–Einstein statistics and Gibrat's Law

We use the familiar model of placing r objects called "stars" in n cells arranged in linear order as in |***|*| |**|, two adjacent bars defining a cell. Let r_k be the number of stars in the kth cell from the left, and let $R = (r_1, r_2, \cdots, r_n)$ be a vector indicating a particular assignment of r stars in n cells where $r = \sum_{k=1}^{n} r_k$. [In the above example, $R = (3, 1, 0, 2)$.] Bose–Einstein statistics assume that each star is indistinguishable from each other star, hence two such arrangements $R = (r_1, r_2, \cdots, r_n)$ and $R' = (r_1', r_2', \cdots, r_n')$ with $\sum_{k=1}^{n} r_k = \sum_{k=1}^{n} r_k' = r$ are said to be indistinguishable if and only if $r_k = r_k'$ for all $k = 1, 2, \cdots, n$ and said to be distinguishable otherwise.

Then, Bose–Einstein statistics postulate that each distinguishable arrangement of r stars in n cells has an equal probability of occurrence. (For example, 2-0, 1-1, 0-2, each has probability 1/3 instead of 1/4, 1/2, 1/4 as under familiar Maxwell–Boltzman statistics.)

There are $(n + r - 1)!/(n - 1)!r!$ distinguishable arrangements of r stars in n cells, since this is the number of ways of arrainging the $(n - 1)$ partitioning bars among r stars, with two additional fixed bars at the end of the array to define the end cells (p. 38 and following pages of ref. 1). Hence, the probability of obtaining a given R is $(n - 1)!r!/(n + r - 1)!$.

We shall first show that Bose–Einstein statistics can be obtained from the Gibrat's law of proportionality applied to the process of throwing in stars when the total number of cells is fixed at n. Let the size of the kth cell, s_k, be the number of stars in this cell plus one. Thus, the size may be interpreted as the number of spaces in the cell between two stars or between two bars or between a star and a bar. Then, the aggregate size of n cells

$$s = \sum_{k=1}^{n} s_k = n + r. \qquad [1]$$

We require under the Gibrat's law of proportionality that the probability that the $r + 1$st star will fall in the kth cell be equal to s_k/s.

Let

$$S = (s_1, s_2, \cdots, s_n) \qquad [2]$$

be a vector of sizes of n cells after r stars have been thrown in. Let

$$S_k = (s_1, s_2, \cdots, s_k - 1, \cdots, s_n), \qquad [3]$$

namely S with the kth component of S reduced by one. S can be obtained from S_k by getting the rth star in the kth cell, providing $s_k - 1 > 0$ since, by definition, the size of a cell can never be less than 1. Assume that Bose–Einstein statistics hold for arrangements of $r - 1$ stars in n cells. Then, the probability of obtaining S_k, denoted by $P(S_k)$, is

$$P(S_k) = (n - 1)!(r - 1)!/(n + r - 2)! \quad \text{if } s_k - 1 \geq 1$$
$$= 0 \quad \text{if } s_k - 1 = 0 \qquad [4]$$

The conditional probability of obtaining S given S_k is $(s_k - 1)/(n + r - 1)$. Therefore, the probability of obtaining S, denoted by $P(S)$, is

$$P(S) = \sum_{k=1}^{n} \frac{s_k - 1}{n + r - 1} P(S_k). \qquad [5]$$

However, if $P(S_k) = 0$, then $s_k - 1 = 0$. Hence, $P(S_k)$ in [5] may be replaced by $(n - 1)!(r - 1)!/(n + r - 2)!$ without affecting the value of $P(S)$, i.e.,

$$P(S) = \frac{(n - 1)!(r - 1)!}{(n + r - 2)!} \times \sum_{k=1}^{n} \frac{s_k - 1}{n + r - 1} = \frac{(n - 1)!r!}{(n + r - 1)!} \qquad [6]$$

since $\sum_{k=1}^{n} s_k - 1 = n + r - n = r$. This shows that S is Bose–Einstein if the distribution before throwing in the rth star is Bose–Einstein and the rth star is placed according to the Gibrat's law of proportionality. Since for $r = 0$, each cell is of size 1 and the Bose–Einstein condition is trivially satisfied, the distribution of r stars in n cells is always Bose–Einstein if stars are thrown in according to Gibrat's Law.

Let us now consider a distribution, $f(i,r)$, of the number of cells with size i after r stars have been thrown in. Let $n_1, n_2, \cdots, n_m$ be a sequence of the values of $f(i,r)$, $i = 1, 2, \cdots, r + 1$, omitting those that are equal to 0. Then, there are $n!/n_1!n_2!\cdots n_m!$ arrangements that lead to the given size distribution and each of them has the same probability $(n - 1)!r!/(n + r - 1)!$ of occurrence under Bose–Einstein statistics. Hence, the probability of obtaining a given size distribution, denoted by $P(f)$, is

$$P(f) = \frac{n!}{n_1!n_2!\cdots n_m!} \frac{(n - 1)!r!}{(n + r - 1)!} \qquad [7]$$

Clearly, the probability is maximum when $n_1 = n_2 = \cdots = n_m = 1$, i.e., when no two cells are of equal size. Also, it is obvious that there is no steady state distribution. However, steady state distributions can be obtained if the number of cells n is also allowed to increase proportionately as the number of stars is increased. Depending upon how new cells are created, we derive two distinct distributions which we shall discuss next.

Two limiting distributions for Bose–Einstein statistics

Let $p(i,s)$ be the probability that a cell will have size i when the aggregate size of all cells is s. Also let $p(i)$ be the steady state probability that a cell will have size i, i.e., $p(i) = \lim_{s\to\infty} p(i,s)$.

Under certain boundary conditions, Gibrat's Law is known to produce as its limiting distribution the Pareto Law, given by:

$$p(i) = Ki^{-\rho} \text{ in which } K, \rho \text{ are constant parameters.} \qquad [8]$$

Under other boundary conditions, the limiting distribution for Bose–Einstein statistics is the geometric distribution:

$$p(i) = M\rho^{-i} \text{ in which } M, \rho \text{ are constant parameters } (\rho > 1). \qquad [9]$$

Eqs. [8] and [9] can be derived by considering the necessary conditions for a steady state of a stochastic process based on Bose–Einstein statistics.

The Pareto distribution

Consider, first, a process in which not only stars but also bars are added. At each round, either a bar or a star is selected with probability α and $1 - \alpha$, respectively. If a star is selected, it is thrown in according to Bose–Einstein statistics, so that each space has an equal chance of receiving it. If a bar is selected, however, it is placed next to an existing bar. That is to say, new cells are added at a rate α and all new cells are of unit size. The average size of cells is a random variable with mean $1/\alpha$.

Regardless of whether a bar or a star is selected at any given round, the aggregate size s of all cells is increased by one at the end of the round either because the size of one of the cells is increased by one or because a new cell of size 1 is added. Thus, we may use s not only as the aggregate size but also as a counter for the number of rounds.

Let $f(i,s)$ be the expected value of the number of cells with size i when the aggregate size of all cells is s. Then, for $i = 1$ we have

$$f(1,s + 1) - f(1,s) = \alpha - (1 - \alpha)f(1,s)/s, \qquad [10]$$

where α is the probability that $f(1,s)$ is increased by one and $(1 - \alpha)f(1,s)/s$ is the probability that $f(1,s)$ is decreased by one as a result of a star falling in one of the unit-sized cells. At the steady state, for all $i = 1,2,\ldots.$

$$p(i) = f(i,s + 1)/(\alpha(s + 1)) = f(i,s)/(\alpha s) \qquad [11]$$

where αs is the expected value of the total number of cells after s rounds. Setting $i = 1$, we use the right-hand equation of [11] to eliminate $f(i,s + 1)$ from [10]:

$$f(1,s)/s = \alpha - (1 - \alpha)f(1,s)/s \qquad [12]$$

$$f(1,s) = \alpha s/(2 - \alpha) \qquad [13]$$

If we define $\rho = 1/(1 - \alpha)$, then,

$$p(1) = f(1,s)/\alpha s = 1/(2 - \alpha) = \rho/(1 + \rho). \qquad [14]$$

For $i > 1$, we have

$$f(i,s + 1) - f(i,s) = (1 - \alpha)[(i - 1)f(i - 1,s)/s - if(i,s)/s] \qquad [15]$$

Using [11] in [15],

$$f(i,s)/s = (1 - \alpha)(i - 1)f(i - 1,s)/s - (1 - \alpha)if(i,s)/s \qquad [16]$$

$$f(i,s)/f(i - 1,s) = (1 - \alpha)(i - 1)/(1 + (1 - \alpha)i) = (i - 1)/(i + \rho) \qquad [17]$$

Thus, $p(i)/p(i - 1)$ is also equal to $(i - 1)/(i + \rho)$. From this and [14], $\rho(i)$ is uniquely given by,

$$p(i) = \frac{\rho}{1 + \rho} \cdot \frac{1}{2 + \rho} \cdot \frac{2}{3 + \rho} \cdot \ldots\ldots \cdot \frac{i - 1}{i + \rho} \qquad [18]$$

Use of the Gamma function,

$$\Gamma(x) = \int_0^\infty c^{-t}t^{x-1}dt \qquad [19]$$

which has the property that,

$$\Gamma(x) = (x - 1)\,\Gamma(x - 1) \qquad [20]$$

and

$$\Gamma(1) = 1 \qquad [21]$$

and the Beta function,

$$\beta(x,y) = \int_0^1 t^{x-1}(1 - t)^{y-1}dt = \frac{\Gamma(x)\Gamma(y)}{\Gamma(x + y)} \qquad [22]$$

simplifies the expression [18] to:

$$p(i) = \rho\,\frac{\Gamma(i)\Gamma(\rho + 1)}{\Gamma(i + \rho + 1)} = \rho B\,(i,\rho + 1) \qquad [23]$$

This is exactly the expression for the Yule distribution (3). The cumulative distribution function $F(i) = \sum_{j=i}^{\infty} p(j)$ is

$$F(i) = \sum_{j=i}^{\infty} \rho B(j,\rho + 1) = \sum_{j=i}^{\infty} \rho \int_0^1 t^{j-1}(1 - t)^\rho dt$$
$$= \rho \int_0^1 (1 - t)^\rho \left[\sum_{j=i}^{\infty} t^{j-1}\right] dt$$
$$= \rho \int_0^1 (1 - t)^\rho t^{i-1}/(1 - t)dt = \rho B(i,\rho) \qquad [24]$$

As $i \to \infty$, $B(i,\rho) \to \Gamma(\rho)i^{-\rho}$ (p. 58 of ref. 5). Hence,

$$\lim_{i\to\infty} F(i) = \rho\Gamma\,(\rho)i^{-\rho} = \Gamma\,(\rho + 1)^{-\rho} \qquad [25]$$

which is the Pareto distribution.

The geometric distribution

By a slight change in boundary conditions, we can now derive necessary conditions that the steady-state distribution be geometric. Consider a process of adding stars and bars as before with the same probabilities (i.e., probability α of adding a bar and $1 - \alpha$ of adding a star at each round). However, now a bar as well as a star may be thrown in any space, and in such a way that each space has an equal chance of receiving which ever object (a bar or a star) is selected for that round.

Then a new cell of size 1 is created, and the number of such cells is consequently increased if and only if a bar is selected (probability α) and is placed next to an existing bar. Since a cell with size greater than 1 offers two spaces adjacent to an existing bar while a cell with size 1 offers only one, the total number of spaces adjacent to an existing bar is $f(1,s) + 2 \sum_{i=2}^{\infty} f(i,s) = f(1,s) + 2(\alpha s - f(1,s)) = 2\alpha s - f(1,s)$, where αs is the expected value of the total number of cells. On the other hand, the number of cells with size 1 is decreased if a star is selected (probability $1 - \alpha$) and is placed in a cell of size 1 (probability $f(1,s)/s$). Thus,

$$f(1,s + 1) - f(1,s) = \alpha(2\alpha s - f(1,s))/s - (1 - \alpha)f(1,s)/s. \qquad [26]$$

The left-hand side of the equation is, as before, $f(1,s)/s$ using [11], hence

$$f(1,s) = \alpha^2 s \qquad [27]$$

Therefore,

$$p(1) = f(1,s)/\alpha s = \alpha = (\rho - 1)/\rho \qquad [28]$$

For $i > 1$, we have

$$f(i,s + 1) - f(i,s) = (1 - \alpha)[(i - 1)f(i - 1,s)/s - if(i,s)/s] + \alpha\left[\sum_{j>i}^{\infty} (2/j)jf(j,s)/s - ((i - 2)/i)if(i,s)/s\right] \qquad [29]$$

The first term indicates the effect of a star being thrown and the second term the effect of a bar being thrown. In the latter case, the number of cells with size i is increased by 1 if a cell with size $j(j > i$ and $j \neq 2i - 1)$ is selected to receive the bar (probability $jf(j,s)/s$) *and* the bar is placed in the ith space from the left bar or from the right bar of the cell (probability $2/j$). It is increased by 2 if a cell with size $j = 2i - 1$ is selected and the bar is placed in the ith space from the left and the right bar (probability $1/j$). In either case, the expected value of the increase in $f(i,s)$ if a bar is selected is $\sum_{j>i}^{\infty} 2f(j,s)/s$. On the other hand, the number of cells with size i is decreased by 1 if a cell with size i is selected (probability $if(i,s)/s$) and the bar is placed in any space in the cell other than the two spaces at its ends (probability $(i - 2)/i$). This explains the second term on the right-hand side.

The left-hand side of [29] is $f(i,s)/s$, as before, using [11]. Substituting this in [29] and multiplying both sides of the equation by s,

$$f(i,s) = (1 - \alpha)(i - 1)f(i - 1,s) - (1 - \alpha)if(i,s) + 2\alpha \sum_{j>i}^{\infty} f(j,s) - \alpha(i - 2)f(i,s) \qquad [30]$$

or

$$(1 + i)f(i,s) - (1 - \alpha)(i - 1)f(i - 1,s) = 2\alpha \sum_{j=1}^{\infty} f(j,s). \quad [31]$$

It is then easy to verify that

$$f(i,s) = \alpha^2 s(1 - \alpha)^{i-1} \qquad i = 1, 2, \cdots. \quad [32]$$

is a solution to [31] with the initial condition [27]. Furthermore, [31] may be written as

$$(1 + i)f(i,s) = (1 - \alpha)(i - 1)f(i - 1,s) + 2\alpha(\alpha s - \sum_{j=1}^{i-1} f(j,s)) \quad [33]$$

which indicates that $f(i,s)$ is uniquely determined if $f(j,s)$ is unique for each $j = 1, 2, \cdots, i - 1$. This result together with [27] assures that [32] is the unique solution to [31] with the initial condition [27].

Since $p(i) = f(i,s)/\alpha s$, [32] may be written as

$$p(i) = \alpha(1 - \alpha)^{i-1} = (\rho - 1)\rho^{-i} \quad [34]$$

$$F(i) = (1 - \alpha)^{i-1} = \rho^{-(i-1)} \quad [35]$$

We have now shown that by employing Bose–Einstein statistics with the creation of new cells (throwing a bar) with probability α, we obtain either the Pareto or the geometric as the steady-state distribution, depending upon whether or not a new cell is restricted to be of unit size (a bar in the former case being placed only in a space adjacent to an existing bar).

Interpretation of the stochastic processes

Comparing the two processes described by Eqs. [15] and [29], respectively, we see that they have identical terms for the addition of new stars to the cells, but that [29], leading to the geometric distribution, has two terms describing the splitting of cells that are absent from [15], leading to the Pareto distribution. These additional terms account for the less skewed shape of the former distribution as compared with the latter, for they amount to a death process that increases frequencies for small i and decreases frequencies for large i.

The simpler process of [15] is easy to interpret as an explanation of city size distributions. It postulates that the population increase, through the net excess of births over deaths and through migration from rural or foreign areas is porportional to current city size. Since these are plausible assumptions under many conditions, it is not surprising that the observed distributions often fit the Pareto Law.

The additional terms in the process of [29] have no easy interpretation in terms of processes of city growth. Cities do not usually split; although, rarely, a standard metropolitan area, as defined by the U.S. Census, will be divided into two such areas. But certainly all possible splits—into various pairs of equal and unequal fragments—do not occur with equal frequency. Thus, the derivation of the geometric distribution from Bose–Einstein statistics, although usual in applications in physics, seems not to be relevant to city sizes,* nor does the geometric distribution appear to fit the observed data.

We are grateful to Daniel R. Vining, Jr. for valuable comments on an earlier draft of this paper. This work was supported by a research grant from the National Science Foundation.

1. Feller, W. (1968) *An Introduction to Probability Theory and Its Applications* (John Wiley & Sons, New York), 3rd ed.
2. Hill, B. M. (1974) "The rank-frequency form of Zipf's Law," *J. Amer. Statist. Ass.* **69**, 1017–1026.
3. Simon, H. A. (1955) "On a class of skew distribution functions," *Biometrika* **42**, 425–440.
4. Simon, H. A. (1960) "Some further notes on a class skew distribution functions," *Inf. Control* **3**, 80–88.
5. Titchmarsh, E. C. (1939) *The Theory of Functions* (Oxford University Press, Fairlawn, N.J.), 2nd Ed.
6. Haran, E. G. P. & Vining, Daniel R., Jr. (1973) "On the implications of a stationary urban population for the size distribution of cities," *Geogr. Anal.* **5**, 296–308.

* This final statement needs to be interpreted carefully. It does *not* assert that the geometric distribution may not be derived from other models approximately obeying Gibrat's Law. Indeed Haran and Vining (6) have published such a derivation, whose assumptions can be given a reasonable interpretation in terms of migration processes. Their derivation, however, is not based on Bose–Einstein statistics.